JHEP 03 (2023) 033

2211.02058

Statistically optimal observables for global SMEFT fits

LHC EFT WG (Area 3)

24/10/23

Jaco ter Hoeve

VU Amsterdam & Nikhef theory group

in collaboration with R. G. Ambrosio, M. Madigan, J. Rojo and V.Sanz

Where theory and experiment meet

We are progressively moving through the simulation chain (latent space)

$$p(x|c) \sim \int dz_{\mathrm{det}} dz_{\mathrm{shower}} dz_{\mathrm{parton}} p(x|z_{\mathrm{det}}) p(z_{\mathrm{det}}|z_{\mathrm{shower}}) p(z_{\mathrm{shower}}|z_{\mathrm{parton}}) p(z_{\mathrm{parton}}|c)$$

DELPHES fast simulation

MADGRAPH5 aMC@NLO

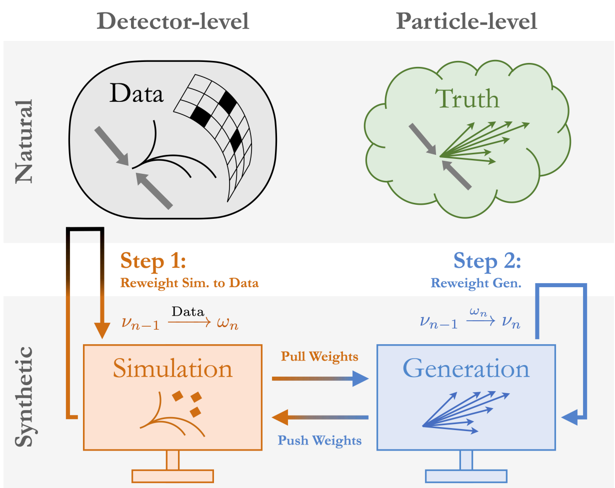
SMEFT@NLO

(Likelihood free) Inference	Observed	→	Hidden/latent variables	→	POI
Generation	Observed	←	Hidden/latent variables	←	POI

Where theory and experiment meet

We are progressively moving through the simulation chain (latent space)

$$p(x|c) \sim \int dz_{\text{det}} dz_{\text{shower}} dz_{\text{parton}} p(x|z_{\text{det}}) p(z_{\text{det}}|z_{\text{shower}}) p(z_{\text{shower}}|z_{\text{parton}}) p(z_{\text{parton}}|c)$$



Unbinned unfolding, Omnifold [1911.09107]

SMEFT@NLO

(Likelihood free) Inference	Observed	→	Hidden/latent variables	→	POI
Generation	Observed	←	Hidden/latent variables	←	POI

Likelihood free inference

- Starting from two balanced datasets $\mathscr{D}_{\mathrm{SM}}$ and $\mathscr{D}_{\mathrm{EFT}}$ drawn from $f(\boldsymbol{x}\,|\,\mathrm{SM})$ and $f(\boldsymbol{x}\,|\,\mathrm{EFT})$, we minimise e.g. the cross-entropy loss

$$L[g(\boldsymbol{x})] = -\frac{1}{N}\sum_{e\in\mathscr{D}_{\mathrm{EFT}}} w_e \log(1-g(\boldsymbol{x}_e)) - \frac{1}{N}\sum_{\mathscr{D}_{\mathrm{SM}}} w_e \log g(\boldsymbol{x}_e)$$

Event weights

$\{m_{t\bar{t}}, \eta_l, \Delta\phi, \ldots\}$

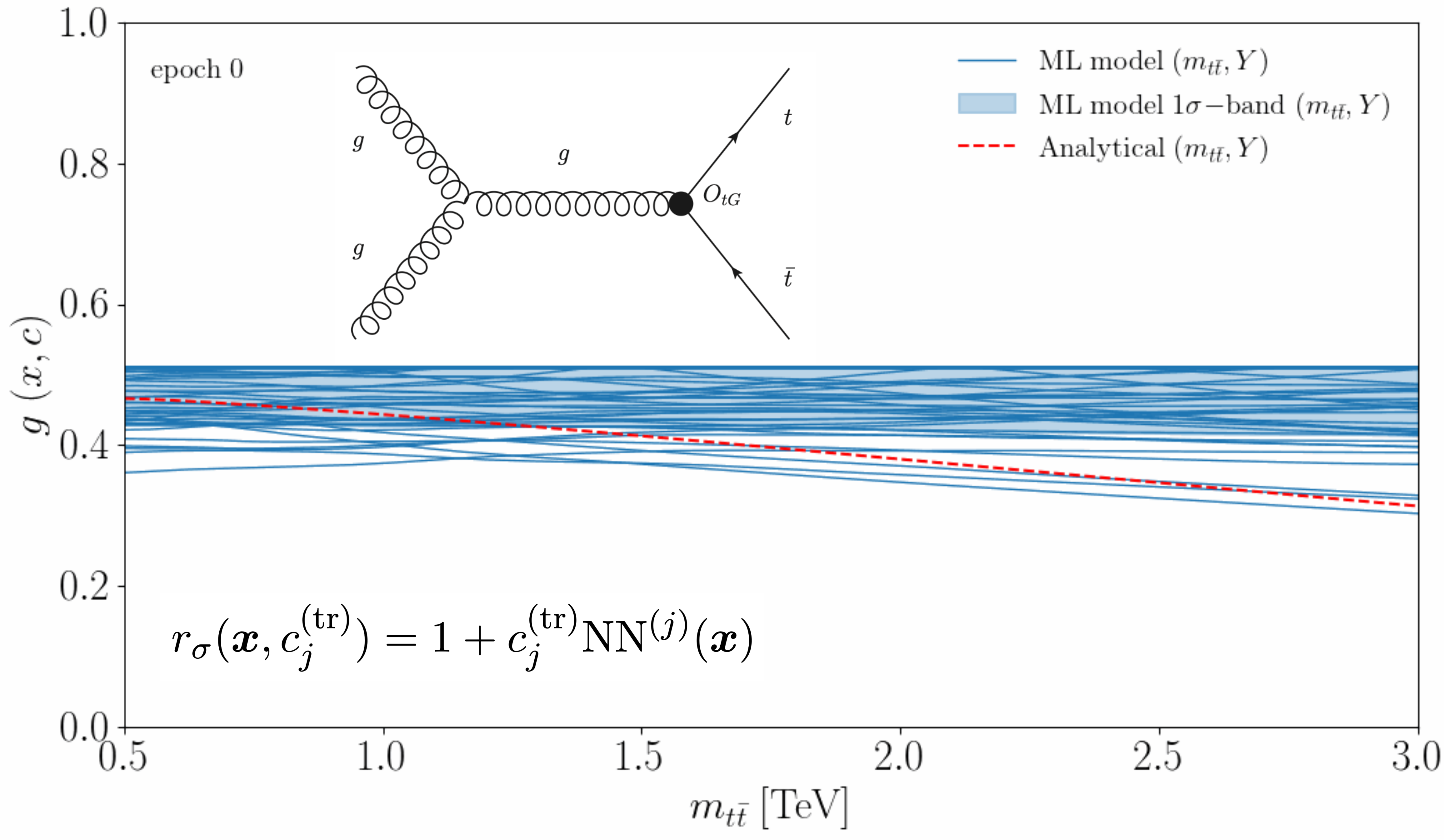
- The learned decision boundary $g(\boldsymbol{x})$ is one-to-one with the likelihood ratio (LR) as $N \to \infty$

$$\frac{\delta L}{\delta g} = 0 \implies \hat{g}(\boldsymbol{x}) = \left(1 + \frac{f(\boldsymbol{x}\,|\,\mathrm{EFT})}{f(\boldsymbol{x}\,|\,\mathrm{SM})}\right)^{-1} \equiv \frac{1}{1+r(\boldsymbol{x})}$$

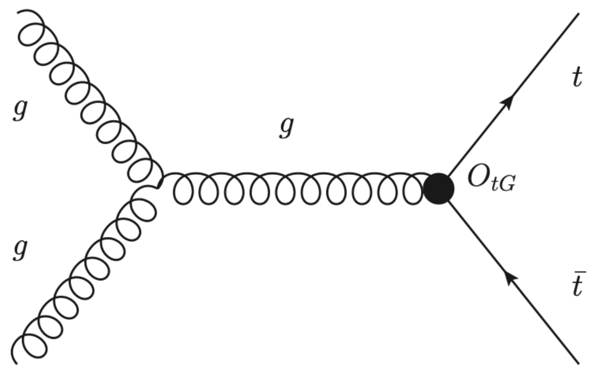
Parameterise with NNs

Likelihood free inference

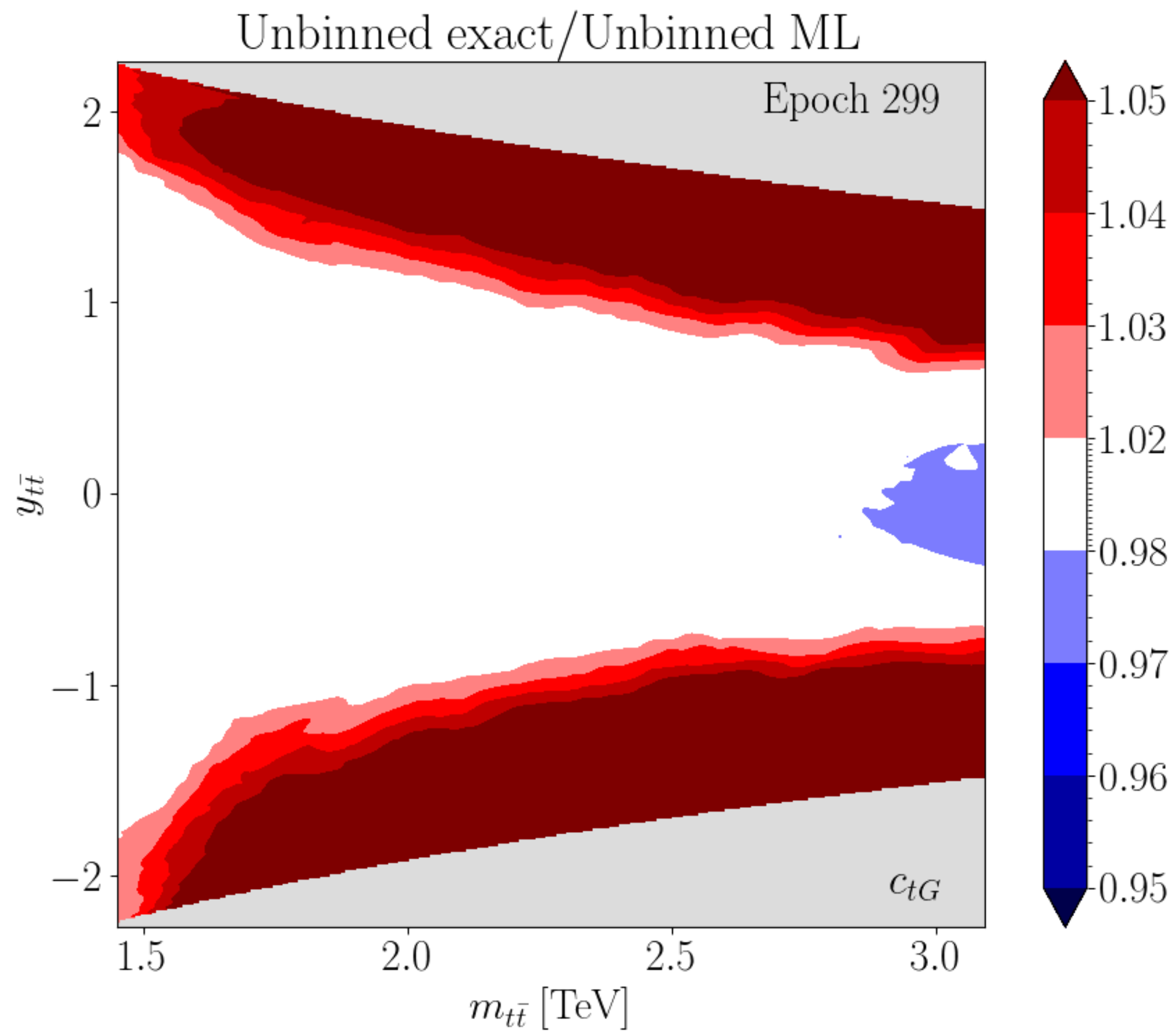
$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\frac{1}{N}\sum_{e\in\mathcal{D}_{\rm eft}} w_e \log(1-g(\boldsymbol{x}_e,\boldsymbol{c})) - \frac{1}{N}\sum_{e\in\mathcal{D}_{\rm sm}} w_e \log g(\boldsymbol{x}_e,\boldsymbol{c})$$

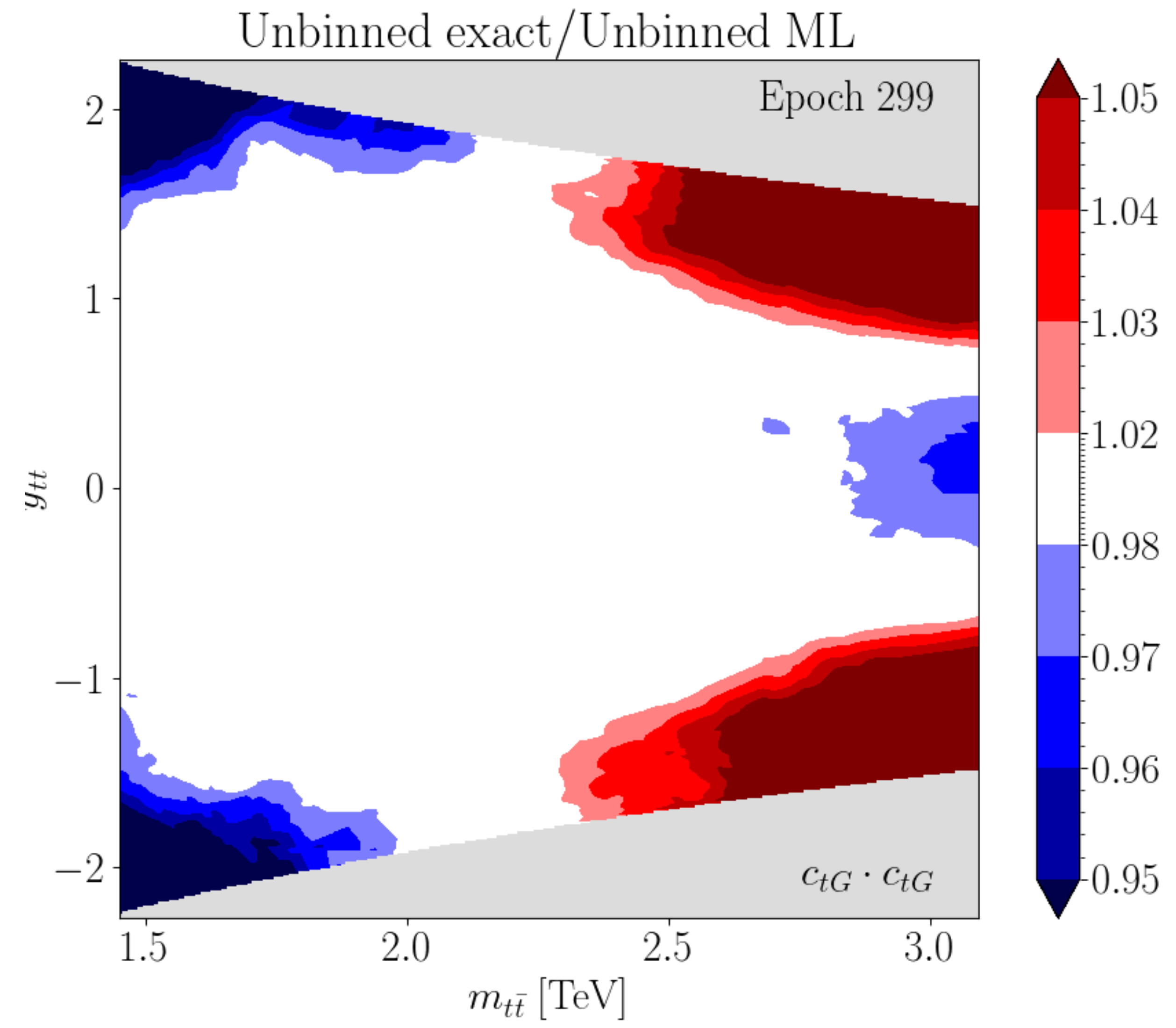


Likelihood free inference

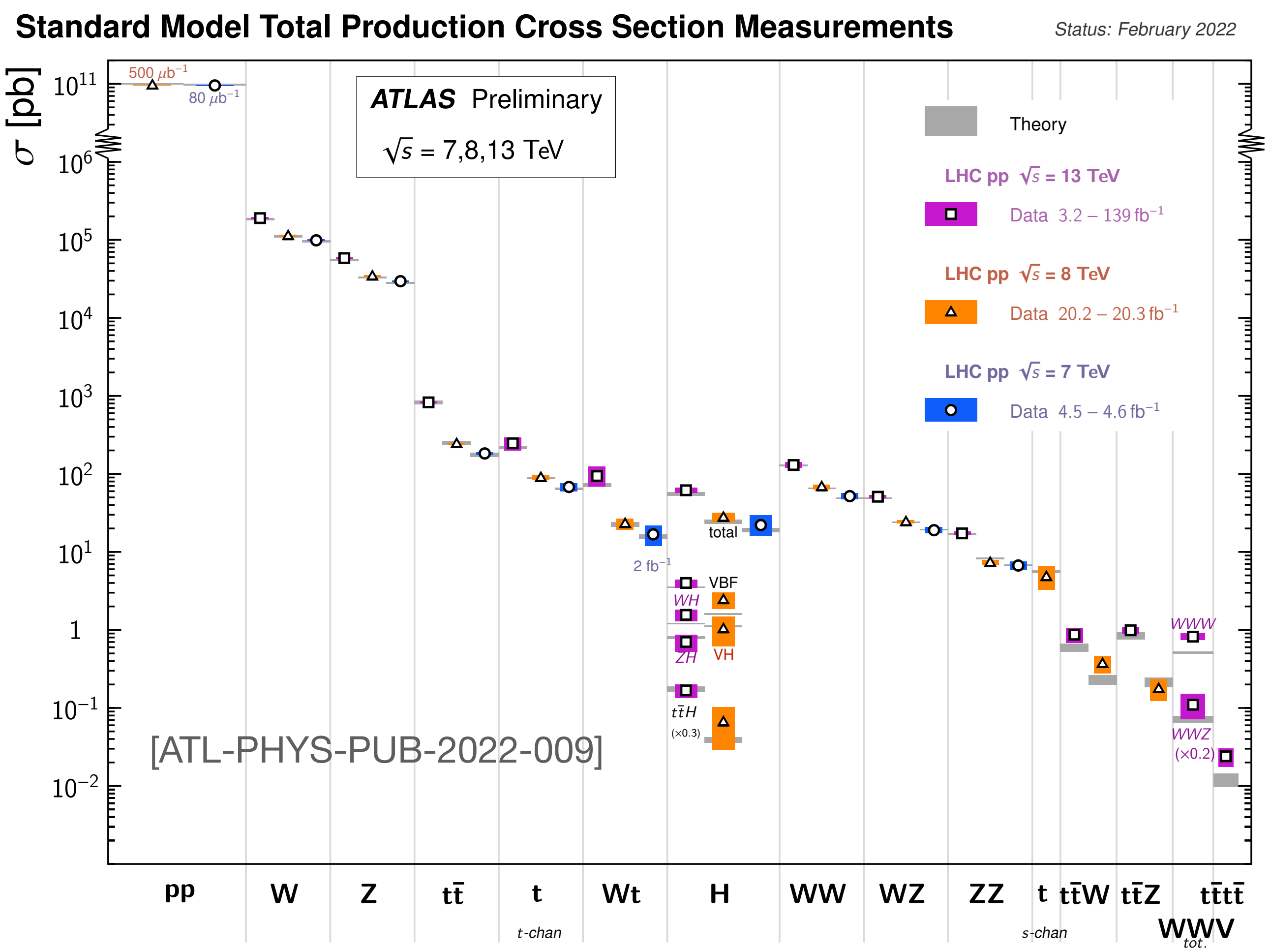


2 kinematic features



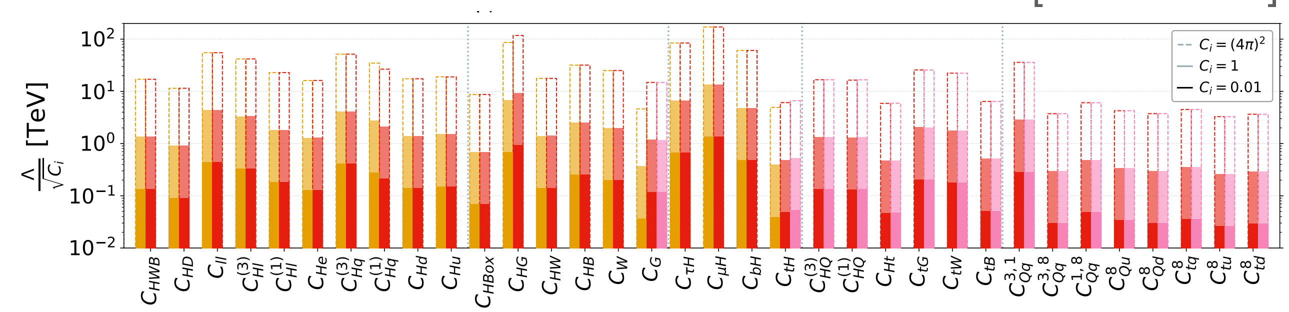


Why interesting for the SMEFT?

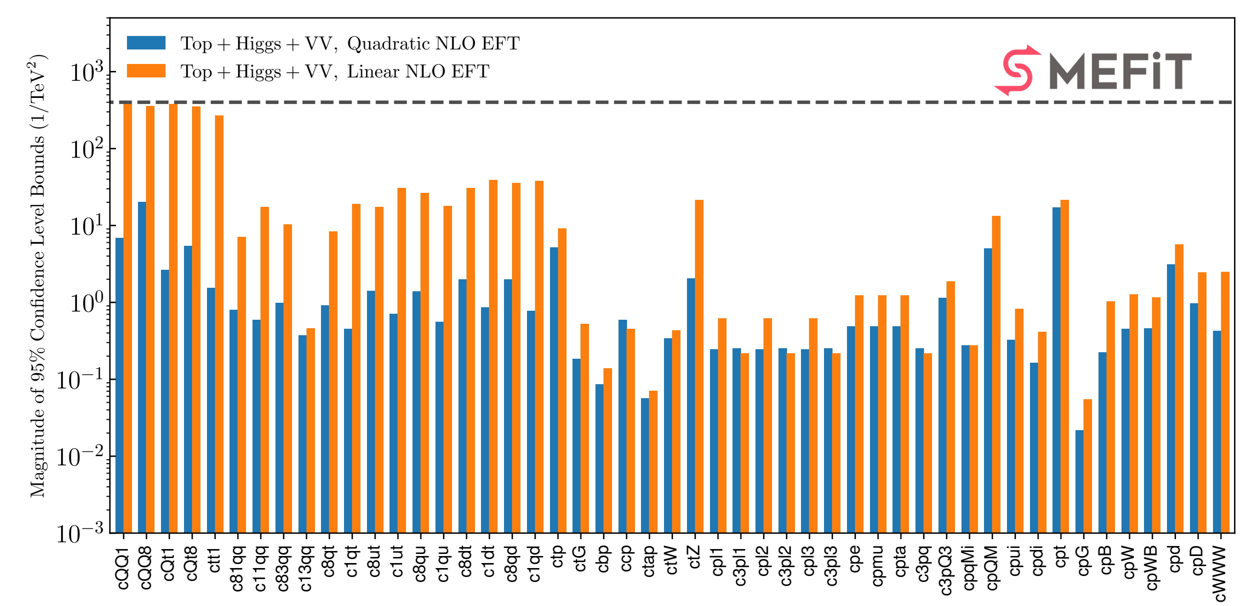


$$\chi^2 \sim \left(\sigma_i(c) - \sigma_{i,\exp}\right)\left(\mathrm{cov}^{-1}\right)_{ij}\left(\sigma_j(c) - \sigma_{j,\exp}\right)$$

[2012.02779]

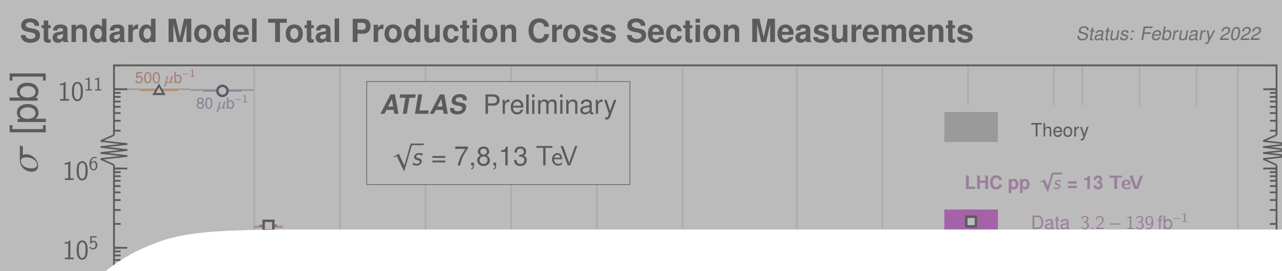


[2302.0660]

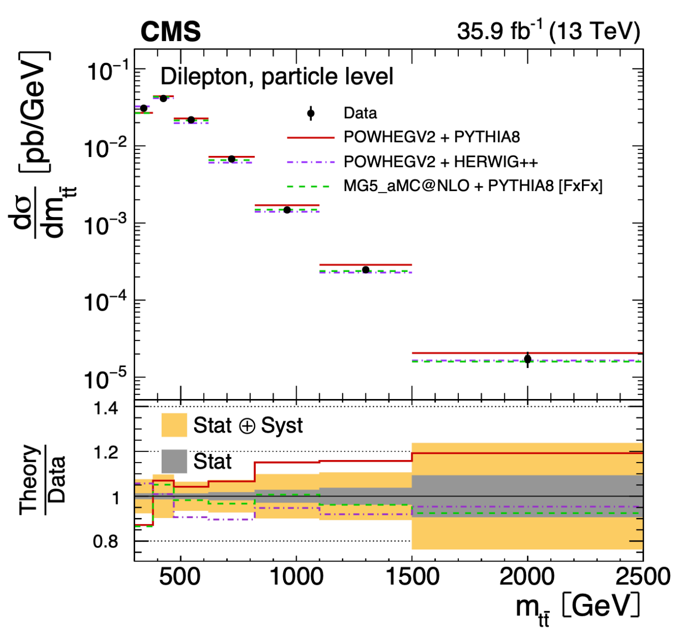


$$\sigma(c) = \sigma_{\mathrm{SM}}\left(1 + \sum_{i}^{N_{d6}} \kappa_i c_i + \sum_{i<j}^{N_{d6}} \tilde{\kappa}_{ij} c_i \cdot c_j\right)$$

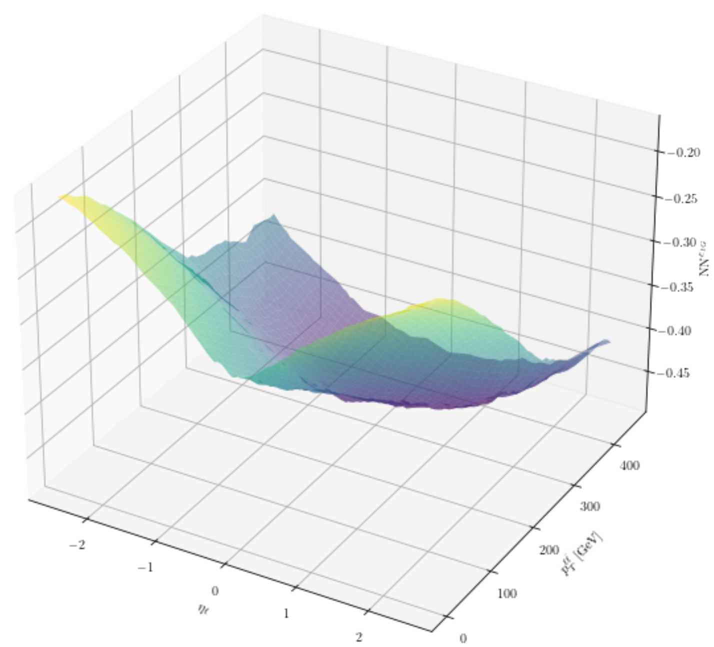
Why interesting for the SMEFT?



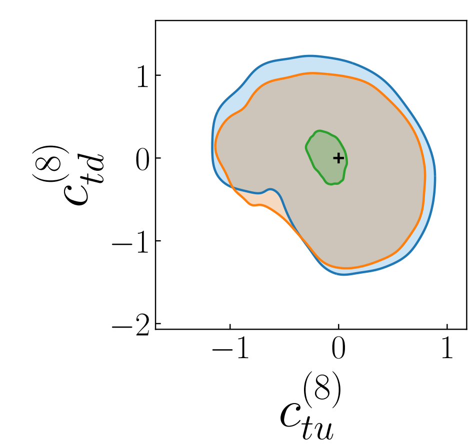
$$\chi^2 \sim \left(\sigma_i(c) - \sigma_{i,\mathrm{exp}}\right)\left(\mathrm{cov}^{-1}\right)_{ij}\left(\sigma_j(c) - \sigma_{j,\mathrm{exp}}\right)$$



Binned, univariate

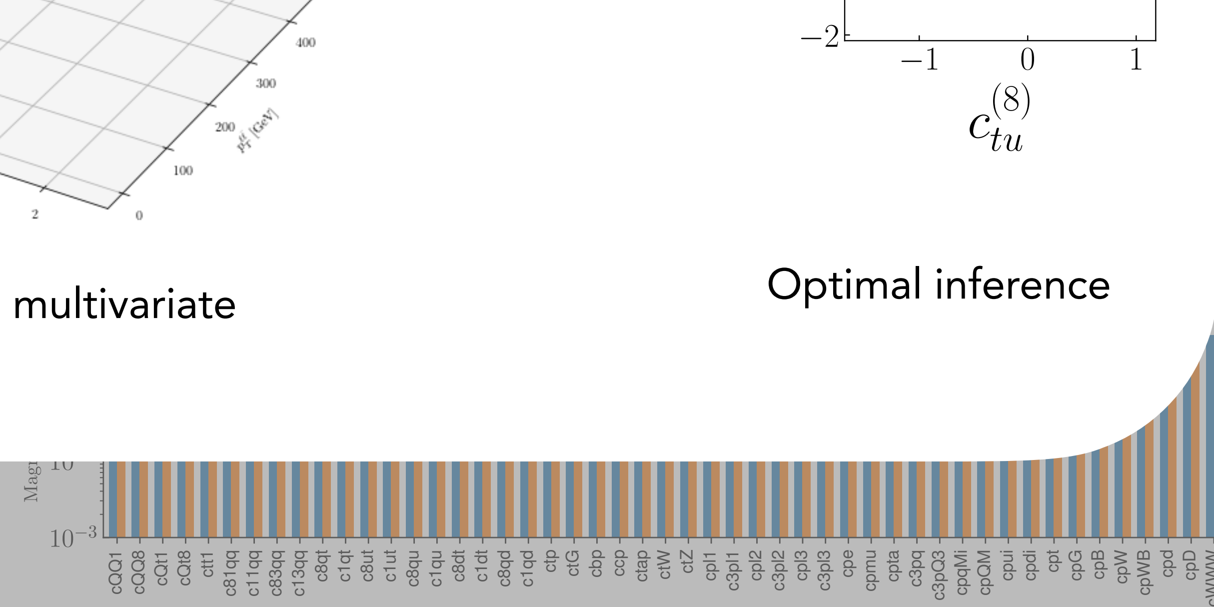


Unbinned, multivariate



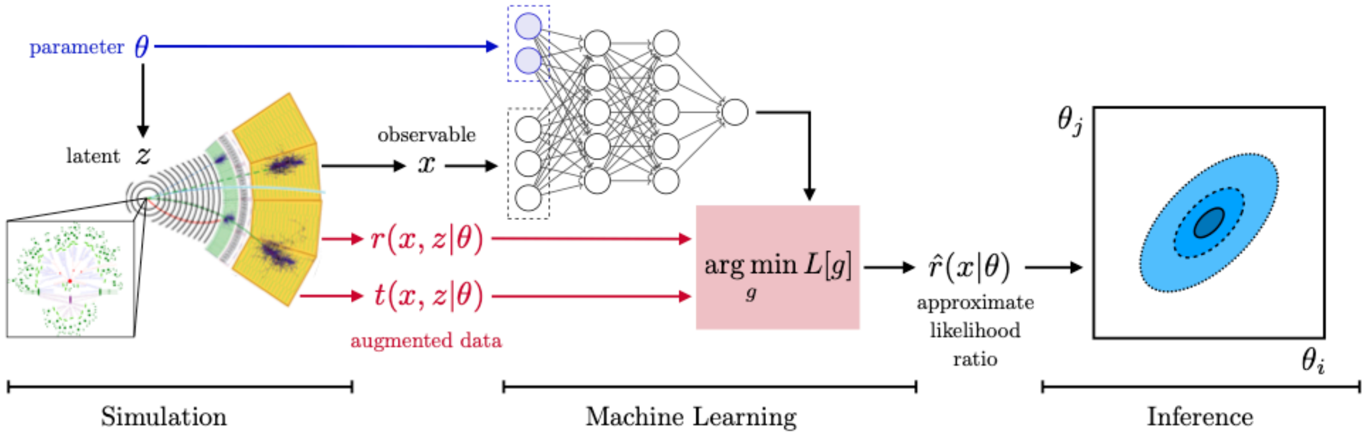
Optimal inference

$$\sigma(c) = \sigma_{\mathrm{SM}}\left(1 + \sum_i \kappa_i c_i + \sum_{i<j} \kappa_{ij} c_i \cdot c_j\right)$$

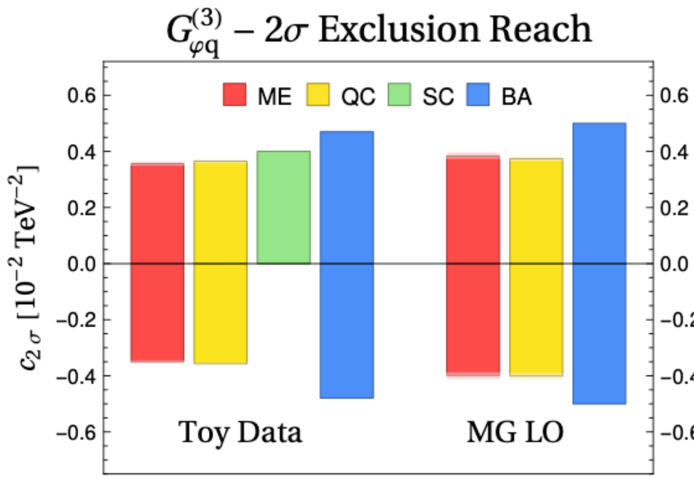


Related work

- Madminer series (J.Brehmer, K.Cranmer, G.Louppe et al.) [1907.10621, 1805.00020, …]
- Parameterized classifiers for SMEFT (A. Glioti et al.) [2007.10356]
- Learning the EFT likelihood with tree boosting (R. Schöfbeck et al) [2205.12976]
- Back to the Formula (A. Butter, T. Plehn et al) [2109.10414]
- Boosted likelihood learning with event reweighing (A. Glioti et al) [2308.05704] See next talk!
- Designing Observables for Measurements with Deep Learning (O.Long, B. Nachman) [2310.08717]



[2010.06439]



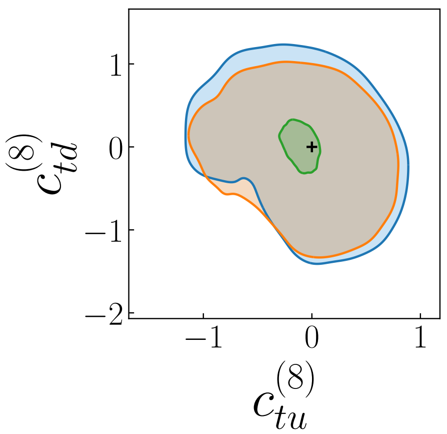
[2007.10356]

This talk

- Global efforts **reinterpret** existing “SM measurements” in an EFT context
- But which measurements are the most **sensitive** to EFT parameters?
  - Inclusive, single to multi-differential (which variables)
  - Binned or unbinned, which binning?

Framework needed to integrate unbinned multivariate observables into **global** SMEFT fits

- **Optimal** bounds on the EFT parameters
- Useful **diagnosis tool** to assess information loss



The ML4EFT framework

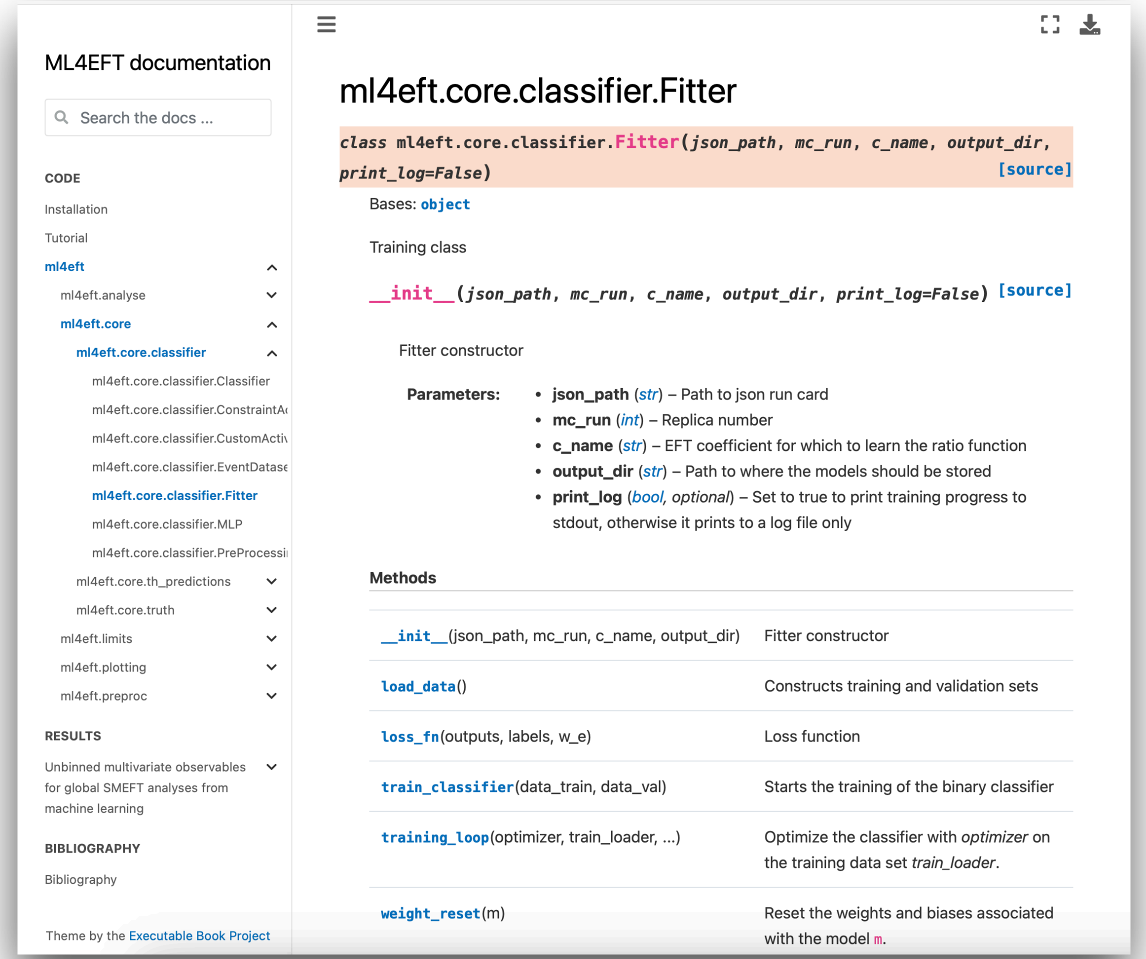
`pip install ml4eft`

https://lhcfitnikhef.github.io/ML4EFT

2211.02058 R. Gomez Ambrosio, JtH, M. Madigan, J. Rojo, V.Sanz

Open-source NN-based python framework for the integration of unbinned multivariate observables into global SMEFT fits

- **Goal**: provide optimal constraints on the SMEFT
- **Diagnostic tool:** what is the information loss incurred by a particular choice of bins?
- **Projections:** how will SMEFT constraints improve if unbinned data are made available?



Modular structure, easy to maintain, well documented

Anticipating global fits

- Global EFT fits typically feature ~50 WCs and thus efficient scaling with the number of WCs becomes essential

- Solution: learn the coefficient functions separately and combine afterwards

$$r(\boldsymbol{x}, \boldsymbol{c}) \equiv \frac{d\sigma(\boldsymbol{x}, \boldsymbol{c})}{d\sigma(\boldsymbol{x}, \boldsymbol{0})} = 1 + \sum_{j=1}^{n_{\rm eft}} r^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k\geq j}^{n_{\rm eft}} r^{(j,k)}(\boldsymbol{x}) c_j c_k$$

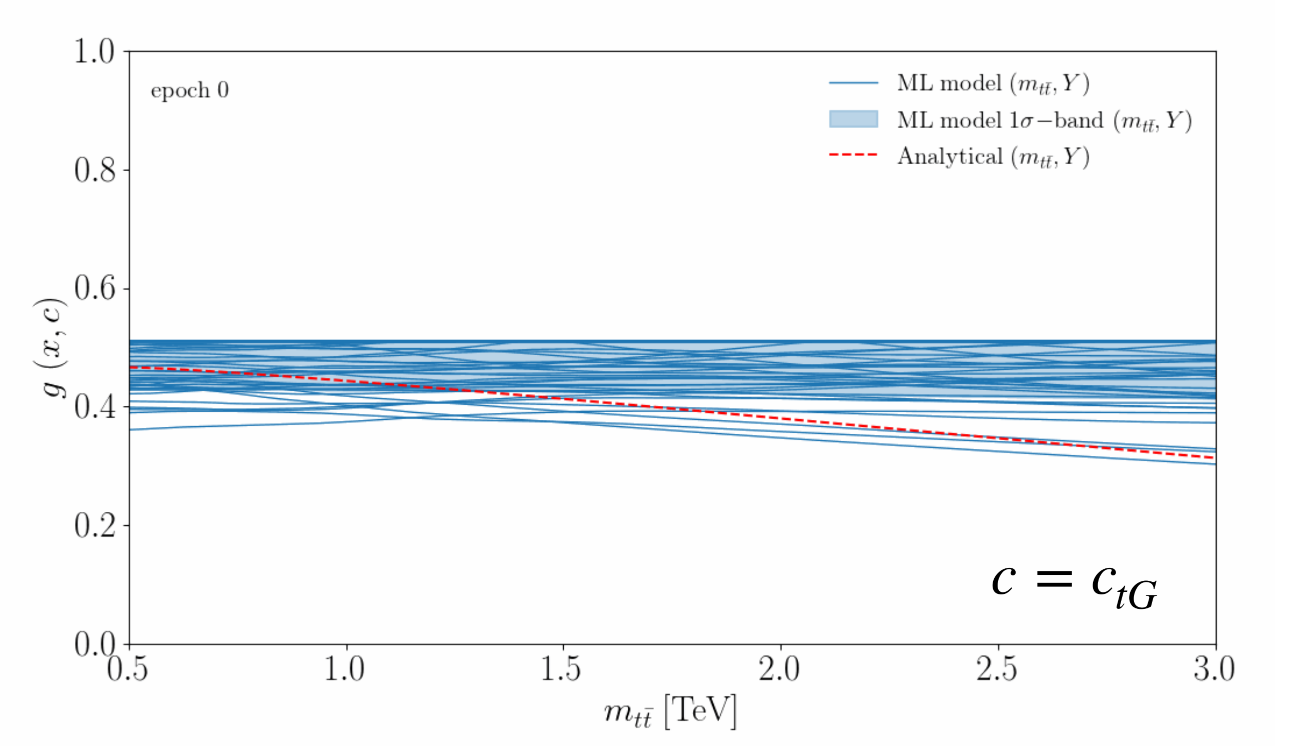
Example: to learn a single $r^{(j)}$, generate $\mathscr{D}_{\rm sm}$ and $\mathscr{D}_{\rm eft}$ at c_j up to $\mathscr{O}(\Lambda^{-2})$. Then $r(\boldsymbol{x}, \boldsymbol{c}) = 1 + r^{(j)}(\boldsymbol{x}) c_j^{(\rm tr)}$ and training means

$$g(\boldsymbol{x}, c_j^{(\rm tr)}) = \left(1 + \left[1 + c_j^{(\rm tr)} \cdot \mathrm{NN}^{(j)}(\boldsymbol{x})\right]\right)^{-1} \qquad \mathrm{NN}^{(j)}(\boldsymbol{x}) \to r^{(j)}(\boldsymbol{x})$$

Uncertainty treatment

- Stick to a regime in which statistical uncertainties dominate over systematics
- Finite training data makes one subject to methodological uncertainties
- Solution: propagate uncertainties to the space of models by training multiple replicas

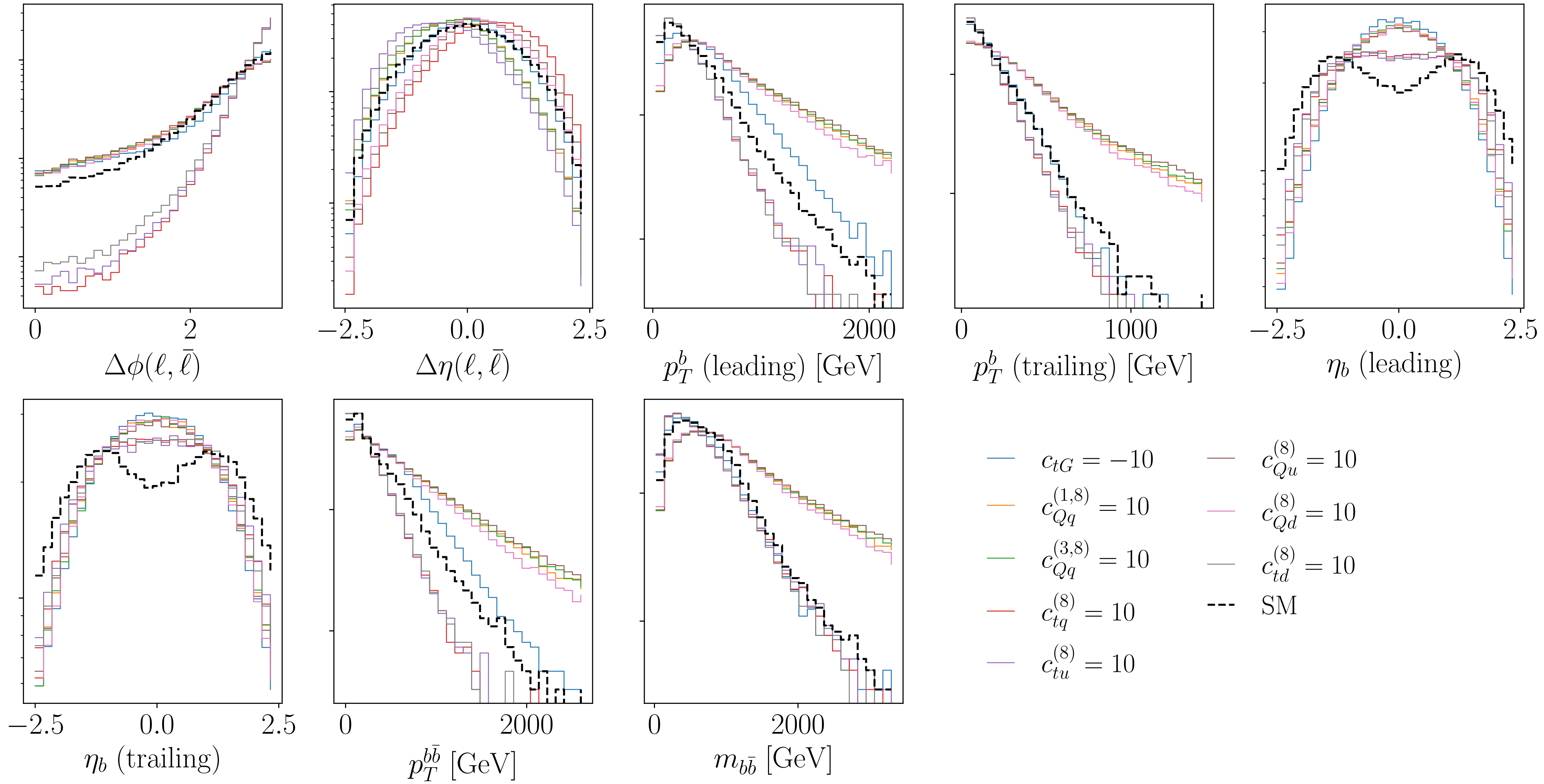
$$\hat{r}^{(i)}(\boldsymbol{x}, \boldsymbol{c}) \equiv 1 + \sum_{j=1}^{n_{\rm eft}} \mathrm{NN}_i^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k \geq j}^{n_{\rm eft}} \mathrm{NN}_i^{(j,k)}(\boldsymbol{x}) c_j c_k \,, \qquad i = 1, \ldots, N_{\rm rep}$$



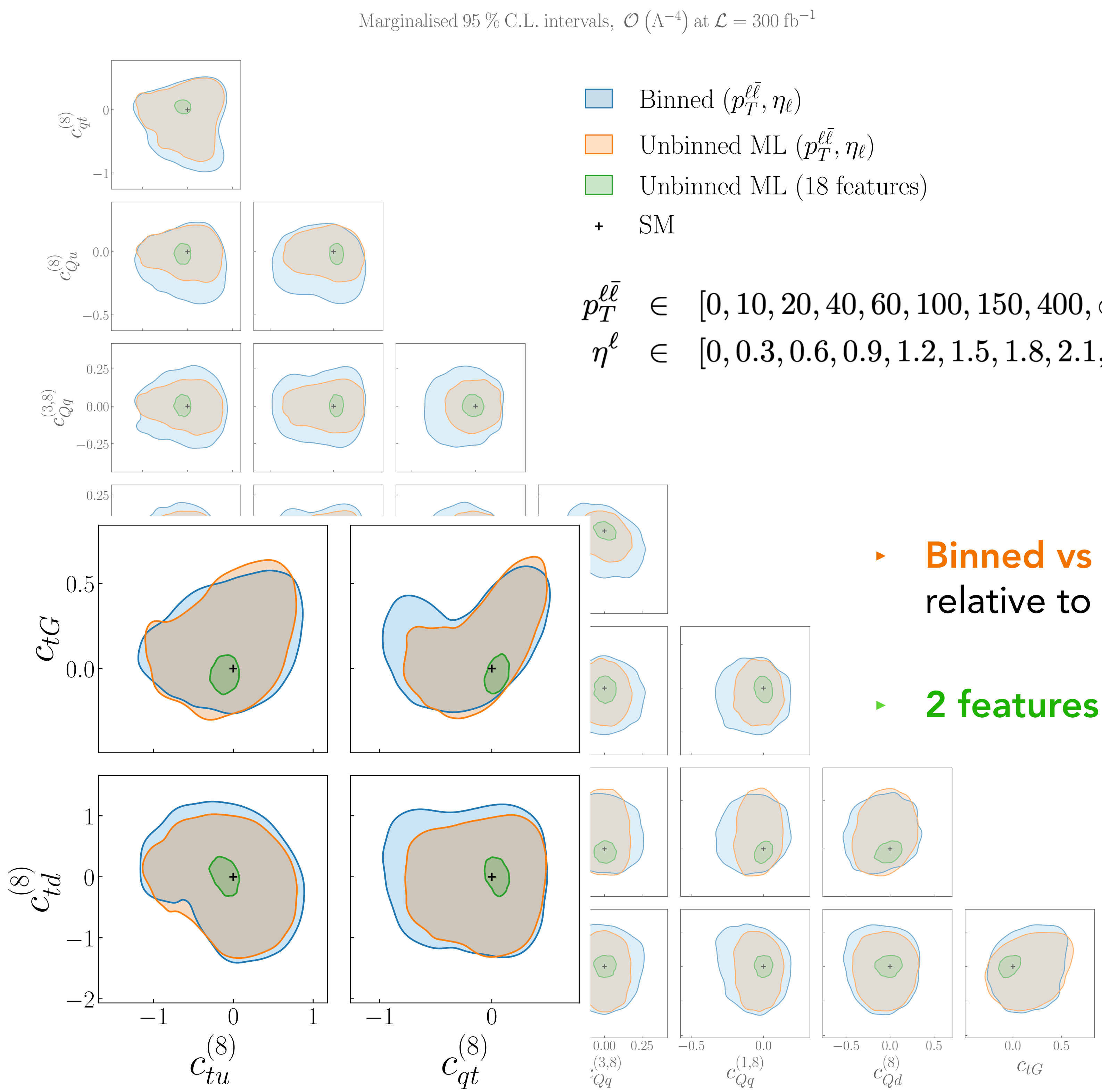
Process	$N_{\rm rep}$	$\widetilde{N}_{\rm ev}$ (per replica)	$N_{\rm nn}$	#trainings
$pp \to t\bar{t}$	50	10^5	4	200
$pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$	25	10^5	40	1000

Let's go multivariate

- $pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$: 18 features, 8 EFT coefficients
- $pp \to hZ \to b\bar{b}\ell^+\ell^-$: 7 features, 7 EFT coefficients



Unbinned observables in the top sector



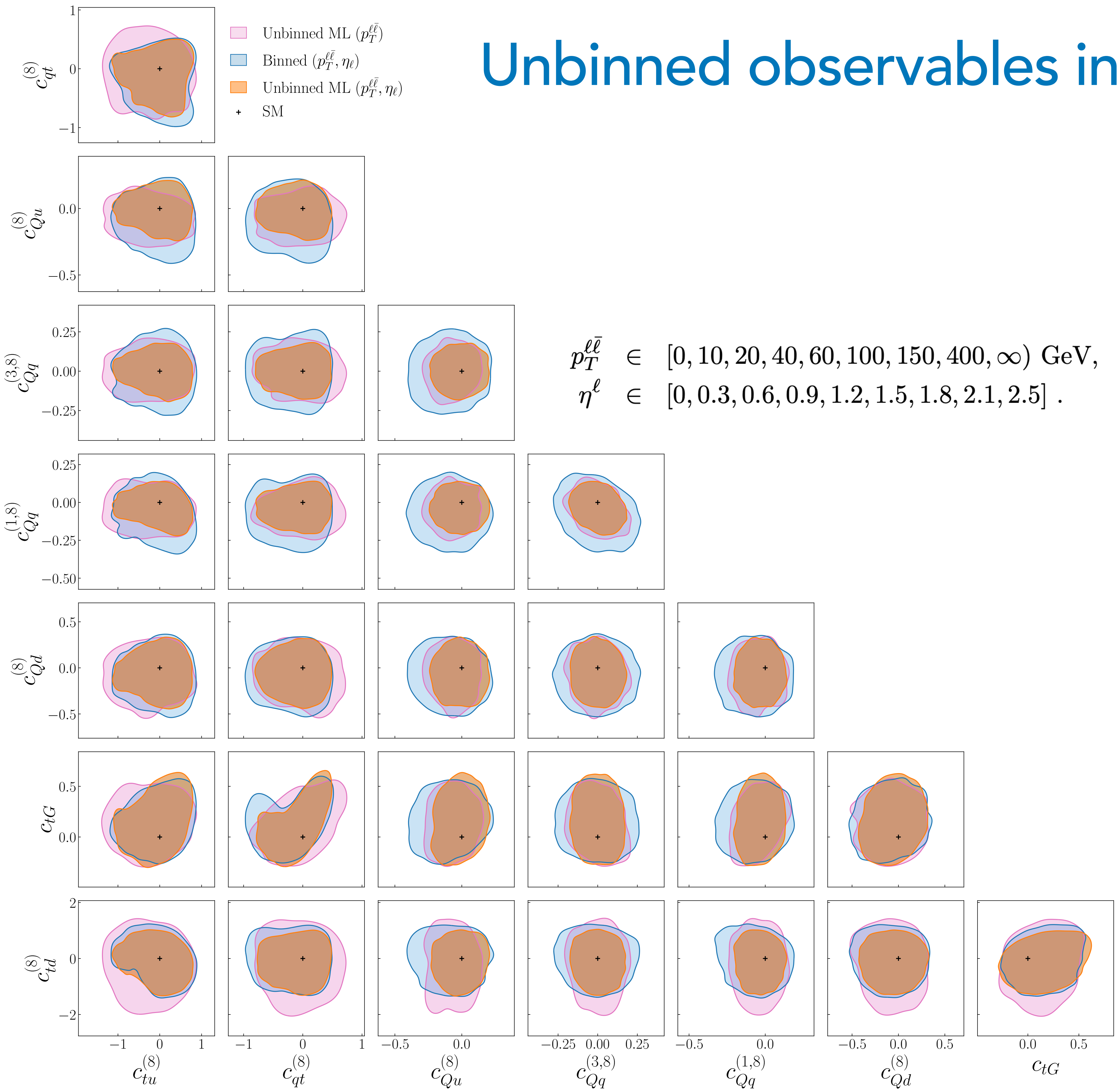
$$p_T^{\ell\bar{\ell}} \in [0, 10, 20, 40, 60, 100, 150, 400, \infty)\ \mathrm{GeV},$$
$$\eta^\ell \in [0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.5]\ .$$

- **Binned vs unbinned** in $(p_T^{\ell\bar{\ell}}, \eta_\ell)$ small improvement relative to binned setup
- **2 features vs 18 features:** big increase in sensitivity

Unbinned observables in the top sector

$$pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$$

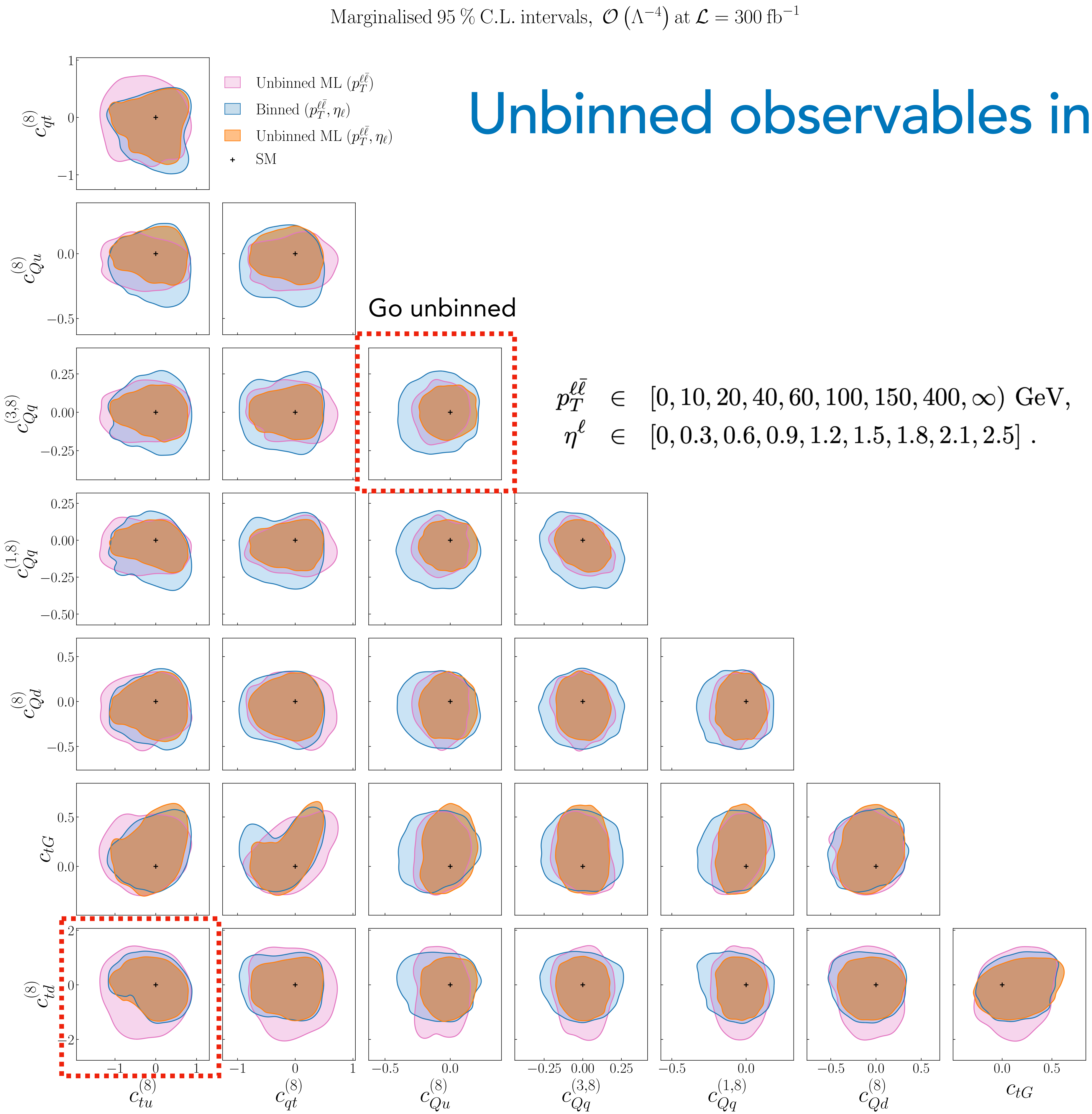
Marginalised 95 % C.L. intervals, $\mathcal{O}\left(\Lambda^{-4}\right)$ at $\mathcal{L} = 300$ fb^{-1}



$$p_T^{\ell\bar{\ell}} \in [0, 10, 20, 40, 60, 100, 150, 400, \infty) \text{ GeV},$$
$$\eta^{\ell} \in [0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.5] \ .$$

- Unbinned multivariate data is advantageous to constrain the EFT parameter space!
- Adding extra features or going unbinned?

Unbinned observables in the top sector

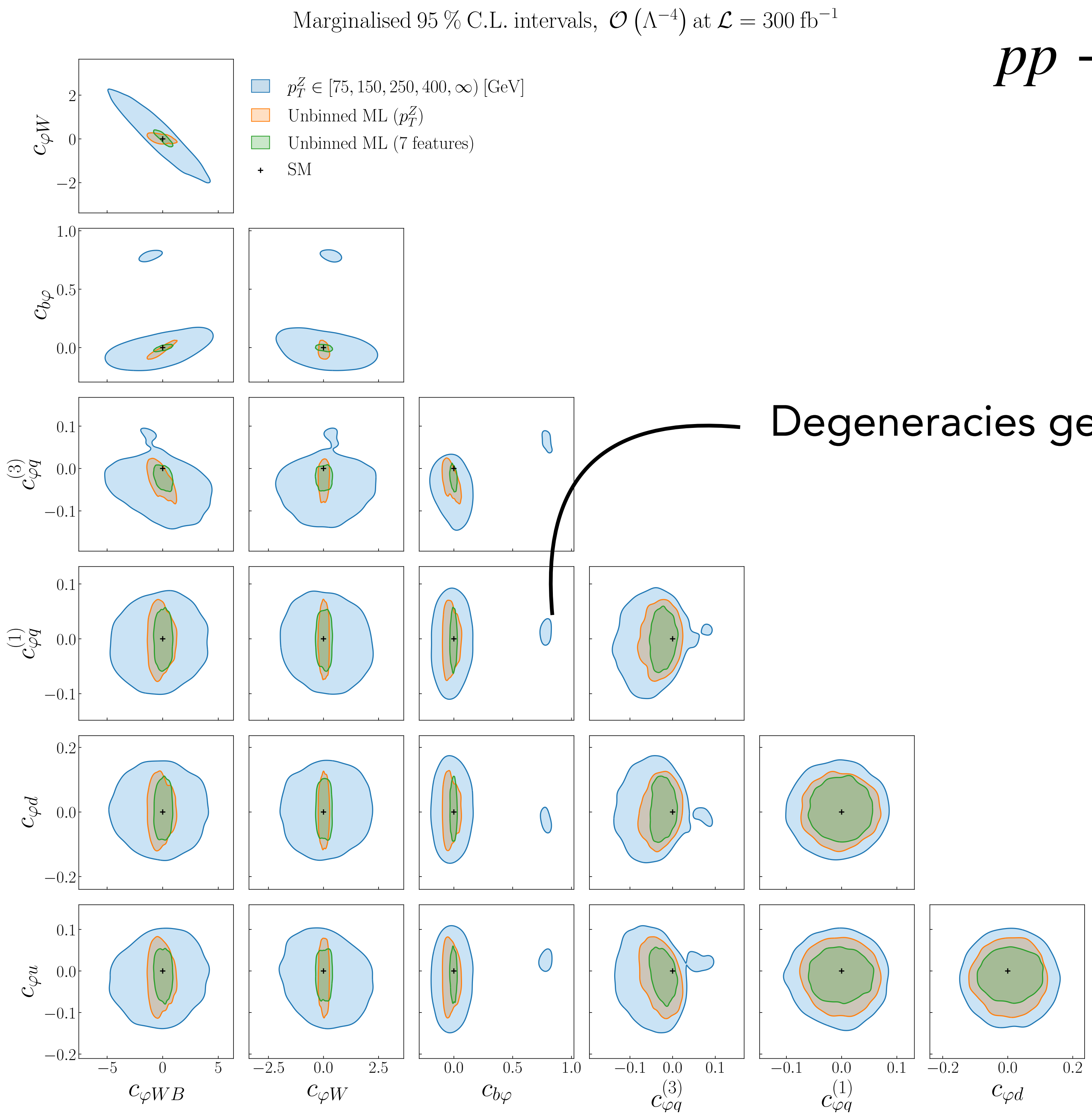


$$pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$$

$$p_T^{\ell\bar{\ell}} \in [0, 10, 20, 40, 60, 100, 150, 400, \infty)\ \mathrm{GeV},$$
$$\eta^\ell \in [0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.5]\ .$$

- Unbinned multivariate data is advantageous to constrain the EFT parameter space!
- Adding extra features or going unbinned?

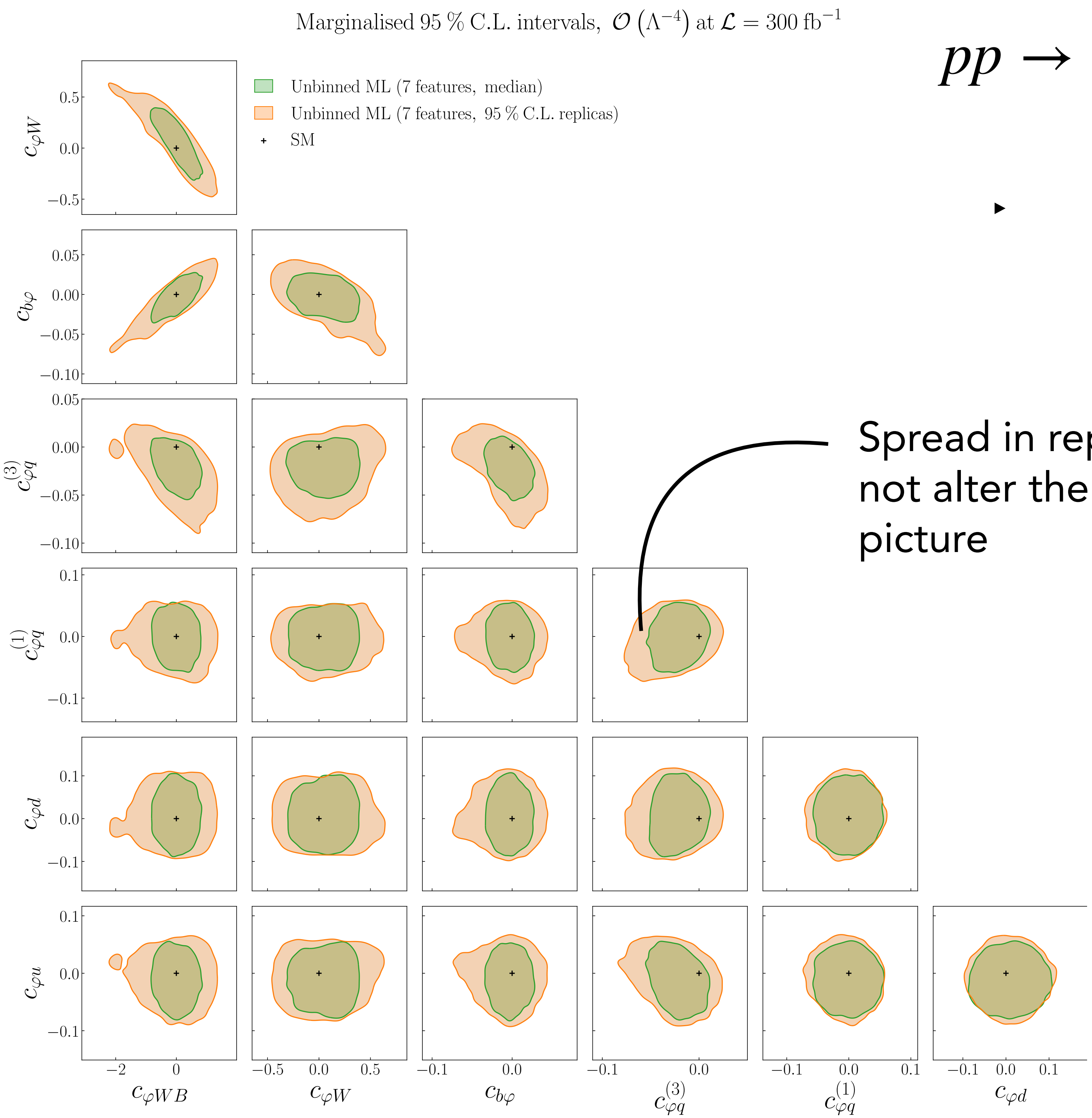
Unbinned observables in Higgs + Z associated production



$$pp \to hZ \to b\bar{b}\ell^+\ell^-$$

Degeneracies get lifted!

Unbinned observables in Higgs + Z associated production

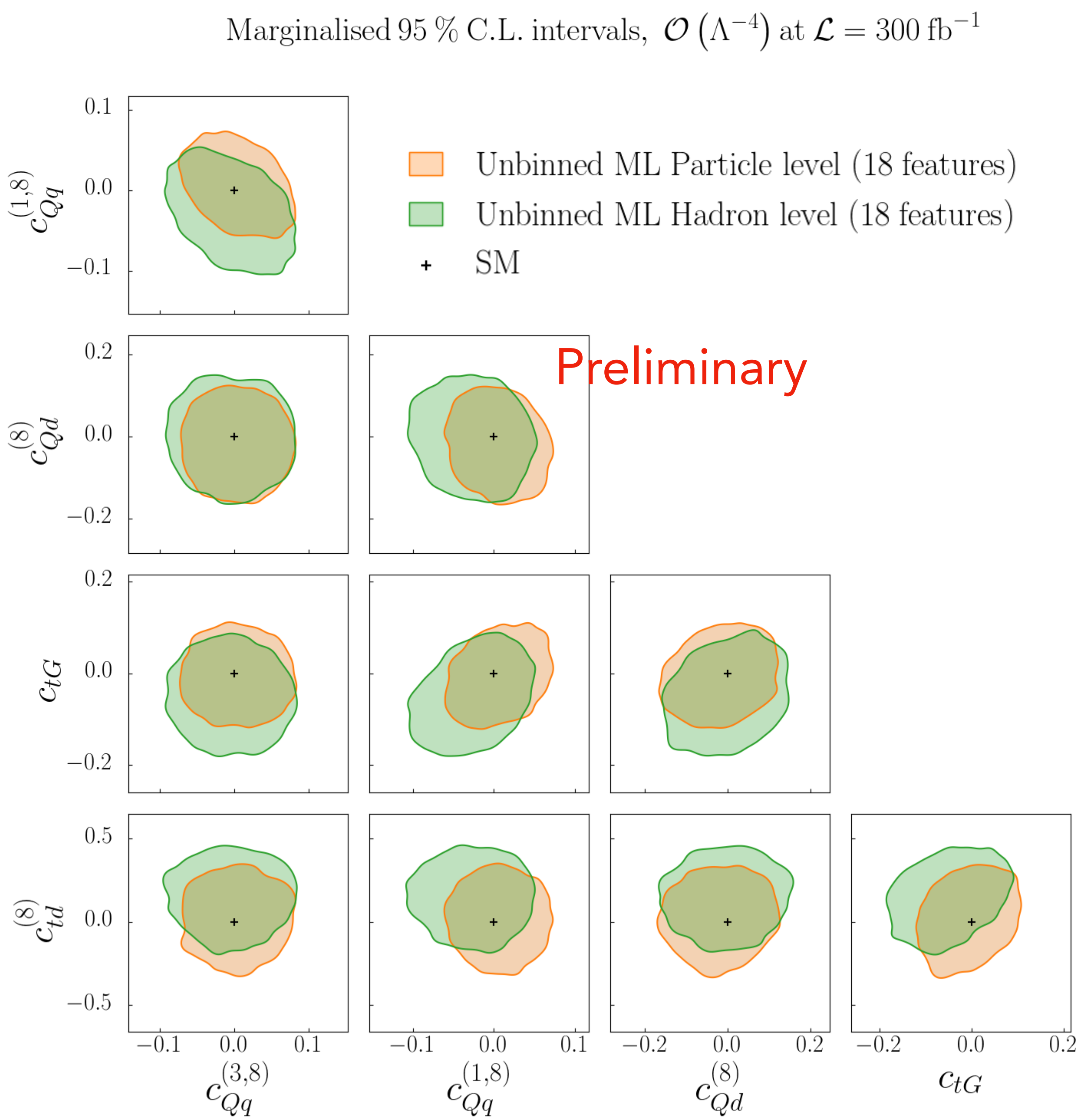


$$pp \to hZ \to b\bar{b}\ell^+\ell^-$$

Spread in replicas does not alter the qualitative picture

Ongoing efforts

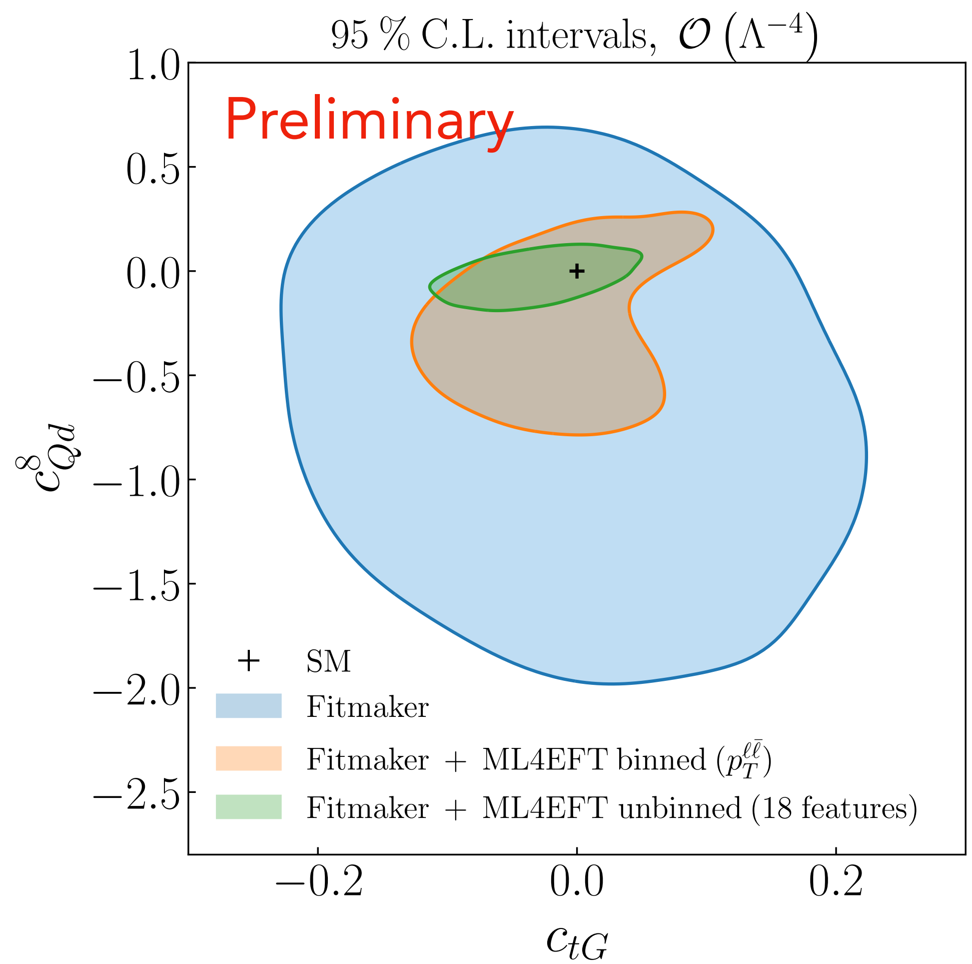
1. Hadronised level



MSc project by Pim Herbschleb

2. Integration into global fits

$$\log \mathcal{L}(c) = \sum_{k=1}^{N_D^{(\text{unbinned})}} \log \mathcal{L}_k^{\text{unbinned}}(c) + \sum_{k=1}^{N_D^{(\text{binned})}} \log \mathcal{L}_k^{\text{binned}}(c)$$



Conclusion and outlook

- Global EFT fits based on unbinned observables **enhance** the sensitivity significantly
- **`ML4EFT`** has efficient scaling properties, as required for global EFT fits and accounts for methodological uncertainties
- WIP: Integration into existing global fit frameworks
- WIP: benchmark study with A. Glioti et al.
- Please visit **`ML4EFT`** on GitHub (documentation + **tutorial**)

lhcfitnikhef.github.io/ML4EFT

Conclusion and outlook

- Global EFT fits based on unbinned observables **enhance** the sensitivity significantly
- **`ML4EFT`** has efficient scaling properties, as required for global EFT fits and accounts for methodological uncertainties
- WIP: Integration into existing global fit frameworks
- WIP: benchmark study with A. Glioti et al.
- Please visit **`ML4EFT`** on GitHub (documentation + **tutorial**)

lhcfitnikhef.github.io/ML4EFT

Thank you!